Moments and distribution of the NPV of a project

Creemers S.

KBI_1623

Moments and distribution of the NPV of a project

Stefan Creemers

IESEG School of Management, Rue de la digue 3, 59000 Lille, France

KU Leuven, Research Center for Operations Management, Naamsestraat 69, 3000 Leuven, Belgium

s.creemers@ieseg.fr

We study the Net Present Value (NPV) of a project with multiple stages that are executed in sequence. A cash flow (positive or negative) may be incurred at the start of each stage, and a payoff is obtained at the end of the project. The duration of a stage is a random variable with a general distribution function. For such projects, we obtain exact, closed-form expressions for the moments of the NPV, and develop a highly accurate closed-form approximation of the NPV distribution itself. In addition, we show two limit theorems that also apply in a more general context (i.e., that also apply for projects where stages are not necessarily executed in sequence). Our work has direct applications in the fields of project selection, project portfolio management, and project valuation. In addition, our work is closely related to the work of CPM/PERT, however, whereas CPM/PERT deals with project completion time, we focus on project NPV.

Keywords: Net Present Value; NPV Distribution; Project Management; Project Evaluation; CPM; PERT

1 Introduction

We consider a project with multiple stages that are executed in sequence, and we assume that the execution sequence is known in advance (i.e., no scheduling decisions are made). Each stage of the project has a random duration with general distribution function. At the start of a stage, a deterministic cash flow (positive or negative) may be incurred, and a deterministic payoff is obtained upon completion of the project. Continuous compounding is used to determine the Net Present Value (NPV) of the project (i.e., the sum of the discounted cash flows that are incurred during the project lifetime). We develop exact, closed-form expressions to obtain the moments of the project NPV distribution. In addition, we provide a highly accurate approximation of the NPV distribution itself. The approximation uses a three-parameter lognormal distribution to match the first three moments of the NPV distribution. A lognormal distribution was chosen because: (1) the moment-matching procedure uses simple, closed-form expressions, and (2) we show that the NPV of a cash flow converges to a lognormal distribution if the cash flow is not incurred during the early stages of the project. We also show that, if a sufficient number of cash flows are incurred, the project NPV converges to a normal distribution. We use examples to illustrate our results, and to show that our approach can easily be implemented, even in a simple spreadsheet application.

Our work has direct applications in the fields of project selection, project portfolio management, and project valuation. In these fields, it is often assumed that: (1) a project is a sequence of stages with cash flows that are incurred at the start of a stage, and (2) a (uncertain) payoff is obtained upon completion of the project (see e.g., Huchzermeier and Loch (2001), Santiago and Vakili (2005), Girotra et al. (2007), and De Reck et al. (2008)). Project selection/investment decisions can be made based on the expected NPV (eNPV) and the risk of a project. Often, the risk of a project is modeled using the variance of the NPV (see e.g., Van Horne (1966)). Other measures of risk are the skewness and/or kurtosis of the NPV, and the probability to have a negative NPV. Higher moments and/or the NPV distribution itself, however, have only been studied using Monte Carlo simulation (see e.g., Savvides (1994) and Kwak and Ingall (2007)). In this article, we develop a closed-form characterization of the NPV distribution of a project that can directly be applied to evaluate project selection/investment decisions.

In the (more operational) field of project scheduling, our work is related to CPM/PERT in the sense that we also focus on a single sequence of stages, and that we also use normal (lognormal) approximations. The study of CPM/PERT dates back to the work of Kelley and Walker (1959) and Malcolmn et al. (1959), and still continues today (refer to Demeulemeester and Herroelen (2002) and Trietsch and Baker (2012) for an overview of the literature). Whereas CPM/PERT deals with the project completion time, we focus on the NPV. In a recent survey, Wiesemann and Kuhn (2015) not only highlight the importance of NPV over project completion time, but also stress the importance of stochastic project scheduling. In stochastic project scheduling, stage durations and/or cash flows are random variables, and as a result, the project NPV is a random variable as well. Almost all of the literature on stochastic project scheduling deals with the maximization of the eNPV (see e.g., Vanhoucke et al. (2001), Szmerekovsky (2005) and Creemers et al. (2010)). In general, it is considered to be impossible to efficiently determine the NPV distribution of a project (Wiesemann and Kuhn 2015). In fact, for the completion time of a project, Hagstrom (1988) has shown that it is $\#P$-complete to determine even a single point of the Cumulative Distribution Function (CDF). Even though the general problem may be impossible to solve, we can still tackle part of it by focussing on a sequence of stages. This is exactly what CPM/PERT does in the context of project completion time, and what we do for the NPV of a project.

The remainder of this article is structured as follows. § 2 develops exact, closed-form expressions for the moments and the distribution of the NPV of a cash flow that is obtained after a single stage. Multiple stages are considered in § 3. In § 3, we also show that the NPV of a single cash flow converges to a lognormal distribution if the cash flow is not incurred during the early stages of the project. § 4 introduces the lognormal approximation of the NPV distribution. In § 5, we develop exact, closed-form expressions for the moments of the NPV distribution of a multi-stage project with intermediate cash flows. In addition, we also show that the NPV of a project converges to a normal distribution, and assess the accuracy of the lognormal and normal approximations of the NPV distribution. § 6 discusses a number of model extensions, and § 7 concludes and provides directions for future research.

2 NPV of a cash flow obtained after a single stage

In this section, we investigate the basic case where a cash flow c is incurred after a single stage. Under continuous compounding, the NPV of a cash flow c is given by:

$$v = ce^{-rt}, \tag{1}$$

where r is the discount rate, and t is the time at which cash flow c is incurred. If t is a realization of T, and if T is a random variable with probability function $f(t)$, the eNPV of the project is given by:

$$\mu = \int_0^\infty f(t)ce^{-rt}\mathrm{d}t, \quad \text{if } T \text{ is continuous,}$$
$$= \sum_{t=0}^{\infty} f(t)ce^{-rt}, \quad \text{if } T \text{ is discrete.} \tag{2}$$

Lemma 1. *Consider a cash flow c that is incurred at time T, where T is a random variable with probability function $f(t)$. Given a discount rate r, the eNPV of c is given by:*

$$\mu = cM_T(-r), \tag{3}$$

where $M_T(u)$ is the Moment Generating Function (MGF) of T.

For notational convenience, let $\phi(r) \equiv M_T(-r)$ such that:

$$\mu = cM_T(-r) = c\phi(r). \tag{4}$$

$\phi(r)$ can be interpreted as the eNPV of a cash flow $c = 1$ that is obtained at time T if discount rate r applies. For most distributions, the MGF (and hence $\phi(r)$) is readily available. There are some distributions, however, for which the MGF does not have a closed-form expression (e.g., the Weibull distribution), or for which the MGF is undefined (e.g., the lognormal distribution). For those distributions, $\phi(r)$ has to be approximated. In addition, note that $\phi(r)$ is not always defined for all values of r. For instance, if T is exponentially distributed, its MGF is given by $M_T(u) = \lambda(\lambda - u)^{-1}$. Hence, if $r = -\lambda$, the MGF about $-r$ is undefined, and μ cannot be determined. In practice, however, this is rarely an issue.

We use an example to illustrate Lemma 1. Consider a cash flow $c = 1,000$ that is incurred at time T, where T follows a gamma distribution with shape parameter $k = 5$ and scale parameter $\tau = 1$. The MGF of the gamma distribution is $M_T(u) = (1 - \tau u)^{-k}$. As a result, $\phi(r) = (1 + \tau r)^{-k}$, and the eNPV of cash flow c is $\mu = c\phi(r) = 620.92$ for discount rate $r = 0.1$.

Theorem 1. *Consider a cash flow c that is incurred at time T, where T is a random variable with probability function $f(t)$. Given a discount rate r, the mean, variance, skewness, and kurtosis of* *the NPV of c are given by:*

$$\begin{aligned}
\mu &= c\phi(r), &(5)\\
\sigma^2 &= c^2(\phi(2r) - \phi^2(r)), &(6)\\
\gamma &= c^3\left(\phi(3r) - 3\phi(2r)\phi(r) + 2\phi^3(r)\right)\sigma^{-3}, &(7)\\
\theta &= \left(\phi(4r) - 4\phi(3r)\phi(r) + 6\phi(2r)\phi^2(r) - 3\phi^4(r)\right)\left(\phi(2r) - \phi^2(r)\right)^{-2}. &(8)
\end{aligned}$$

If we revisit the previous example, the moments of the NPV distribution of cash flow c are: $\mu = 620.92$, $\sigma^2 = 16,334$, $\gamma = -0.2347$, and $\theta = 2.7064$ for discount rate $r = 0.1$.

Theorem 2. *Consider a cash flow c that is incurred at time T, where T is a random variable with probability function $f(t)$. Given a discount rate r, the CDF and Probability Density Function (PDF) of the NPV of cash flow c are given by:*

$$\begin{aligned}
G(v) &= 1 - F\left(\ln\left(\frac{c}{v}\right)r^{-1}\right), &(9)\\
g(v) &= \frac{d}{dv}G(v) = \frac{f\left(\ln\left(\frac{c}{v}\right)r^{-1}\right)}{|r|v}, &(10)
\end{aligned}$$

where $F(t)$ is the CDF of T. Note that: (1) if $r > 0$, then v has range $0 \leq v < c$, (2) if $r = 0$, then $v = c$, and (3) if $r < 0$, then v has range $c < v \leq \infty$.

We illustrate Theorem 2 by means of an example. In the example, a cash flow c is incurred at time T, where T follows an exponential distribution with rate parameter λ. For a given discount rate r, the CDF and PDF of the NPV of cash flow c are:

$$\begin{aligned}
G(v) &= \left(\frac{c}{v}\right)^{-\lambda r^{-1}}, &(11)\\
g(v) &= \frac{\lambda}{|r|v}G(v). &(12)
\end{aligned}$$

Similar, elegant results can easily be obtained for other probability functions.

3 NPV of a cash flow obtained after multiple stages

In this section, we consider the NPV of a cash flow that is incurred after multiple stages. For ease of understanding, we often use payoff p in the explanation of our results (as payoff p is obtained at the end of the project; after all stages have been completed). Note, however, that the results in this section hold for any cash flow that is incurred during the lifetime of the project.

Lemma 2. *Consider a project with multiple stages $w : w \in \mathbf{N} = \{1, 2, \ldots, n\}$ that are executed in sequence. Each stage $w : w \in \mathbf{N}$ has duration distribution $f_w(t)$ and corresponding factor $\phi_w(r)$ that is obtained using (4). If the duration distributions of the individual stages are independent, the duration of the project itself has factor:*

$$\phi_{1,n}(r) = \prod_{w \in \mathbf{N}} \phi_w(r). \tag{13}$$

We can combine Theorem 1 with Lemma 2 to determine the moments of the NPV of a cash flow that is incurred after multiple stages. For instance, consider the NPV of a payoff p that is obtained upon completion of a project with three stages. The stages have factors $\phi_1(r)$, $\phi_2(r)$, and $\phi_3(r)$ respectively. The mean and variance of the NPV of payoff p are given by:

$$\mu = p\phi_1(r)\phi_2(r)\phi_3(r) = p\phi_{1,3}(r), \tag{14}$$

$$\sigma^2 = p^2(\phi_1(2r)\phi_2(2r)\phi_3(2r) - \phi_1^2(r)\phi_2^2(r)\phi_3^2(r)) = p^2(\phi_{1,3}(2r) - \phi_{1,3}^2(r)). \tag{15}$$

The skewness, kurtosis, and higher-order moments can be obtained in the same way.

Lemma 3. *Consider a project with multiple stages $w : w \in \mathbf{N} = \{1, 2, \ldots, n\}$ that are executed in sequence. Each stage $w : w \in \mathbf{N}$ has a duration distribution $f_w(t)$ with mean d_w and variance s_w^2. If the duration distributions of the individual stages are independent, the mean and variance of the project duration are given by:*

$$d_{\mathbf{N}} = \sum_{w \in \mathbf{N}} d_w, \tag{16}$$

$$s_{\mathbf{N}}^2 = \sum_{w \in \mathbf{N}} s_w^2. \tag{17}$$

If n is sufficiently large, the duration of the project will converge to a normal distribution with mean $d_{\mathbf{N}}$ and standard deviation $s_{\mathbf{N}}$.

Lemma 3 is an important and well-known result in the CPM/PERT literature (see e.g., Malcolm et al. (1959), Van Slyke (1963), and Moder and Phillips (1970)), and allows to make predictions on the completion time of a project. We will use Lemma 3 to show that the NPV of a payoff p converges to a lognormal distribution if n is sufficiently large.

Theorem 3. *Consider a project with multiple stages $w : w \in \mathbf{N} = \{1, 2, \ldots, n\}$ that are executed in sequence. If the duration distributions of the individual stages are independent, and if n is sufficiently large, the NPV of payoff p converges to a lognormal distribution with location parameter $\alpha = \ln(p) - rd_{\mathbf{N}}$ and scale parameter $\beta = rs_{\mathbf{N}}$.*

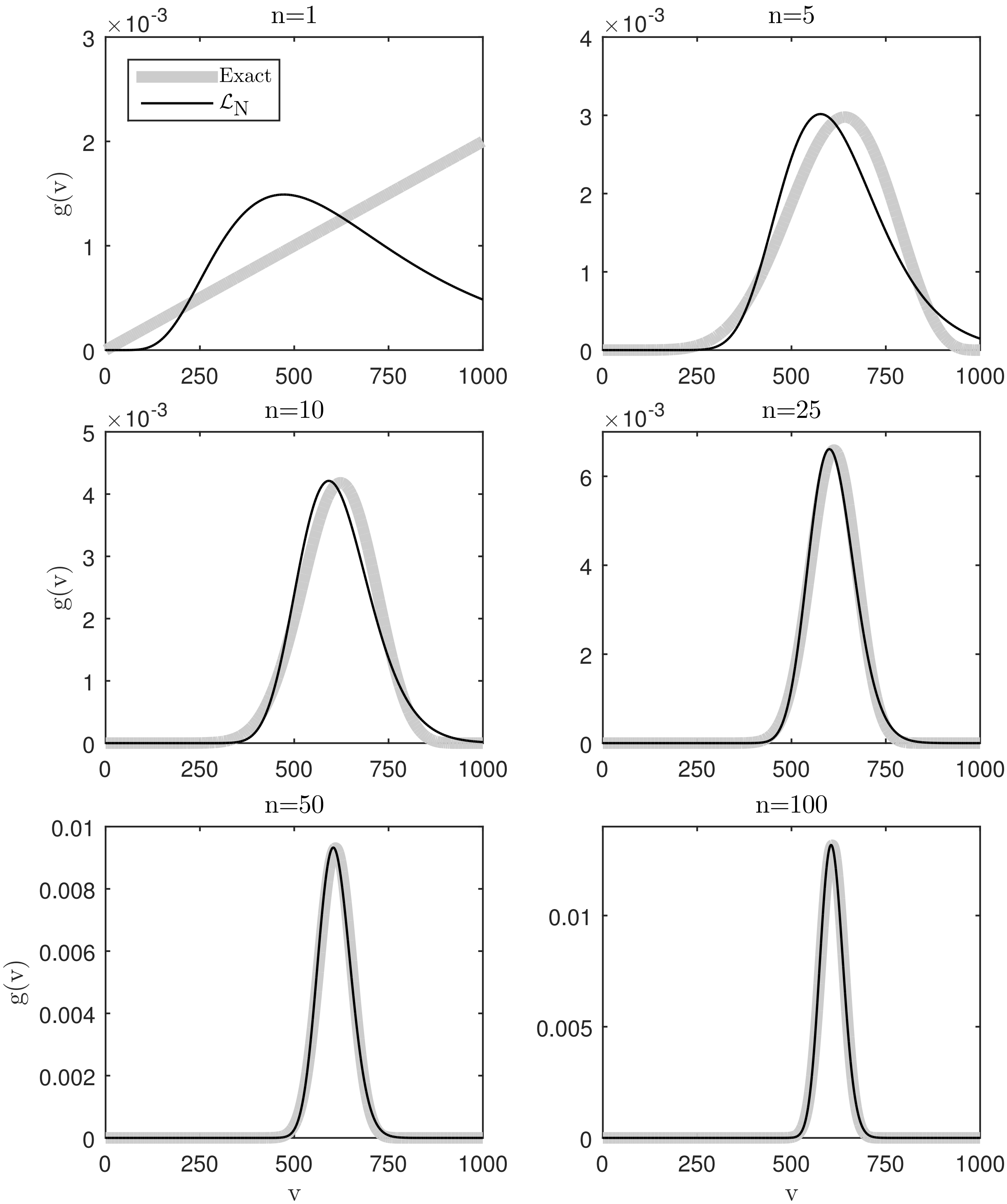
Note that Theorem 3 also applies in a more general context where stages are not necessarily executed in sequence. In fact, as long as a stage w is preceded by a sufficient number of other stages, the NPV of cash flow c_w converges to a lognormal distribution.

In order to illustrate Theorem 3, consider a project with n stages that are executed in sequence, and that have i.i.d. exponential durations with rate parameter λ (i.e., the project duration follows an Erlang distribution with parameters n and λ). A payoff p is obtained upon completion of the project. After applying Theorem 2, we obtain the PDF of the NPV of payoff p:

$$g(v) = \frac{\lambda \left(\frac{p}{v}\right)^{-\lambda r^{-1}} \left(\ln\left(\frac{p}{v}\right) \lambda r^{-1}\right)^{n-1}}{|r| v (n-1)!}. \tag{18}$$

The approximate lognormal distribution has location parameter $\alpha = \ln(p) - rn\lambda^{-1}$ and scale parameter $\beta = r\sqrt{n}\lambda^{-1}$, and is denoted by $\mathcal{L}_{\mathrm{N}}$. Given a payoff $p = 1,000$, and a rate parameter $\lambda = 1$, Figure 1 shows the exact and the approximate PDF of the distribution of the NPV of payoff p for various values of n. The discount rate r is set equal to $0.5n^{-1}$. Table 1 reports the mean, variance, skewness, kurtosis, and Kolmogorov–Smirnov test statistic (i.e., the maximum absolute difference in cumulative probability; the maximum absolute difference between $G(v)$ and the CDF of $\mathcal{L}_{\mathrm{N}}$). We observe that, if n is small, Lemma 3 (and hence Theorem 3) does not hold, and the approximation performs poorly. If, on the other hand, n is large, the approximation is fairly accurate, and the NPV of a payoff p may be approximated by a lognormal distribution.

Figure 1: PDF of the exact NPV and the $\mathcal{L}_N$ approximation for various number of stages



Exact NPV distribution						
n	1	5	10	25	50	100
μ	666.67	620.92	613.91	609.53	608.04	607.29
σ^2	55,556	16,334	8,654	3,589	1,817	914
γ	-0.566	-0.235	-0.163	-0.101	-0.071	-0.050
θ	2.4000	2.7060	2.8300	2.9252	2.9613	2.9803

$\mathcal{L}_{\mathrm{N}}$ approximation						
n	1	5	10	25	50	100
$\mu_{\mathcal{L}\mathrm{N}}$	687.29	621.89	614.16	609.57	608.05	607.29
$\sigma^2_{\mathcal{L}\mathrm{N}}$	134,164	19,829	9,549	3,734	1,853	923
$\gamma_{\mathcal{L}\mathrm{N}}$	1.7500	0.6909	0.4814	0.3018	0.2128	0.1502
$\theta_{\mathcal{L}\mathrm{N}}$	8.8980	3.8606	3.4148	3.1623	3.0806	3.0401
K–S	0.1587	0.0596	0.0421	0.0266	0.0188	0.0133

Table 1: Accuracy of the $\mathcal{L}_{\mathrm{N}}$ approximation for various number of stages

4 A lognormal approximation of the NPV distribution

Theorem 3 only holds for cash flows that are incurred after a sufficient number of stages. Hence, the NPV of a cash flow does not always follow a lognormal distribution. Often, it is impossible to characterize the exact NPV distribution of a cash flow, however, we can use Theorem 1 to obtain its moments. A moment-matching procedure can then be used to define a distribution that approximates the true NPV distribution.

Moment-matching procedures can be evaluated along three lines: (1) the number of moments matched, (2) the computational efficiency, and (3) the generality of the solution. Ideally, a moment-matching procedure uses closed-form expressions to match as many moments as possible under general conditions. Most of the literature on moment matching has focussed on the use of phase-type (PH) distributions (see e.g., Osogami (2005) and Boute et al. (2007)). Using PH distributions, up to three moments can be matched using closed-form expressions (see e.g., Osogami and Harchol-Balter (2006)). In this article, we do not adopt PH distributions, however, use a lognormal approximation of the NPV distribution of a cash flow c. Not only does the lognormal distribution allow us to develop closed-form expressions to match up to three moments of any real-valued distribution with non-zero skew, it is also a logical choice as the NPV distribution of a cash flow c converges to a lognormal distribution if it is incurred after a sufficient number of stages (see also Theorem 3).

In what follows, we define two moment-matching procedures. In a first procedure, we match the first two moments of the NPV distribution. A second procedure matches the first three moments. We use $\mathcal{L}_2$ and $\mathcal{L}_3$ to denote both approximations respectively.

Lemma 4. *We can approximate the NPV distribution by matching its first two moments using a lognormal distribution with scale and location parameter:*

$$\beta = \sqrt{\ln\left(1+\eta^2\right)}, \tag{19}$$

$$\alpha = \ln(\mu) - 0.5\beta^2, \tag{20}$$

where μ and $\eta = \sigma^2\mu^{-2}$ are the mean and Squared Coefficient of Variation (SCV) of the NPV distribution respectively.

In order to match three moments, we use a three-parameter (or bounded) lognormal distribution (see e.g., Aitchison and Brown (1957)) with location, shape, and threshold parameter α, β, and κ respectively. The threshold parameter can be used to bound the support of the distribution, and can either serve as a lower or as an upper bound (for matching distributions with positive/negative skew respectively). The mean, variance, skewness, kurtosis, PDF, and CDF of the three-parameter lognormal distribution are given by:

$$\mu_{\mathcal{L}3} = \kappa + \delta e^{\alpha+0.5\beta^2}, \tag{21}$$

$$\sigma^2_{\mathcal{L}3} = \left(e^{\beta^2} - 1\right) e^{2\alpha+\beta^2}, \tag{22}$$

$$\gamma_{\mathcal{L}3} = \delta\left(2 + e^{\beta^2}\right)\sqrt{e^{\beta^2} - 1}, \tag{23}$$

$$\theta_{\mathcal{L}3} = e^{2\beta^2}\left(3 + e^{\beta^2}\left(2 + e^{\beta^2}\right)\right) - 3, \tag{24}$$

$$g_{\mathcal{L}3}(v) = \frac{1}{\delta(v-\kappa)\beta\sqrt{2\pi}} e^{\left(\frac{(\ln(\delta(v-\kappa))-\alpha)^2}{2\beta^2}\right)}, \tag{25}$$

$$G_{\mathcal{L}3}(v) = \frac{1}{2} - \frac{\delta}{2}\operatorname{Erf}\left(\frac{\alpha - \ln\left(\delta(v-\kappa)\right)}{\beta\sqrt{2}}\right), \tag{26}$$

where $\delta = -1$ if the distribution has negative skew, and $\delta = 1$ otherwise.

Lemma 5. *We can approximate the NPV distribution by matching its first three moments using a*

$\mathcal{L}_2$ approximation						
n	1	5	10	25	50	100
$\mu_{\mathcal{L}2}$	666.67	620.92	613.91	609.53	608.04	607.29
$\sigma^2_{\mathcal{L}2}$	55,556	16,334	8,654	3,589	1,817	914
$\gamma_{\mathcal{L}2}$	1.1049	0.6262	0.4581	0.2958	0.2106	0.1495
$\theta_{\mathcal{L}2}$	5.2463	3.7053	3.3754	3.1560	3.0790	3.0397
K–S	0.1357	0.0597	0.0421	0.0266	0.0188	0.0133

$\mathcal{L}_3$ approximation						
n	1	5	10	25	50	100
$\mu_{\mathcal{L}3}$	666.67	620.92	613.91	609.53	608.04	607.29
$\sigma^2_{\mathcal{L}3}$	55,556	16,334	8,654	3,589	1,817	914
$\gamma_{\mathcal{L}3}$	-0.566	-0.235	-0.163	-0.101	-0.071	-0.050
$\theta_{\mathcal{L}3}$	3.5743	3.0981	3.0470	3.0182	3.0090	3.0045
K–S	0.0590	0.0118	0.0059	0.0023	0.0011	0.0006

Table 2: Accuracy of the $\mathcal{L}_2$ and $\mathcal{L}_3$ approximations for various number of stages

bounded lognormal distribution with parameters:

$$\beta = \sqrt{\ln\left(\frac{2^{1/3}}{\left(2+\gamma^2+\sqrt{4\gamma^2+\gamma^4}\right)^{1/3}}+\frac{\left(2+\gamma^2+\sqrt{4\gamma^2+\gamma^4}\right)^{1/3}}{2^{1/3}}-1\right)}, \tag{27}$$

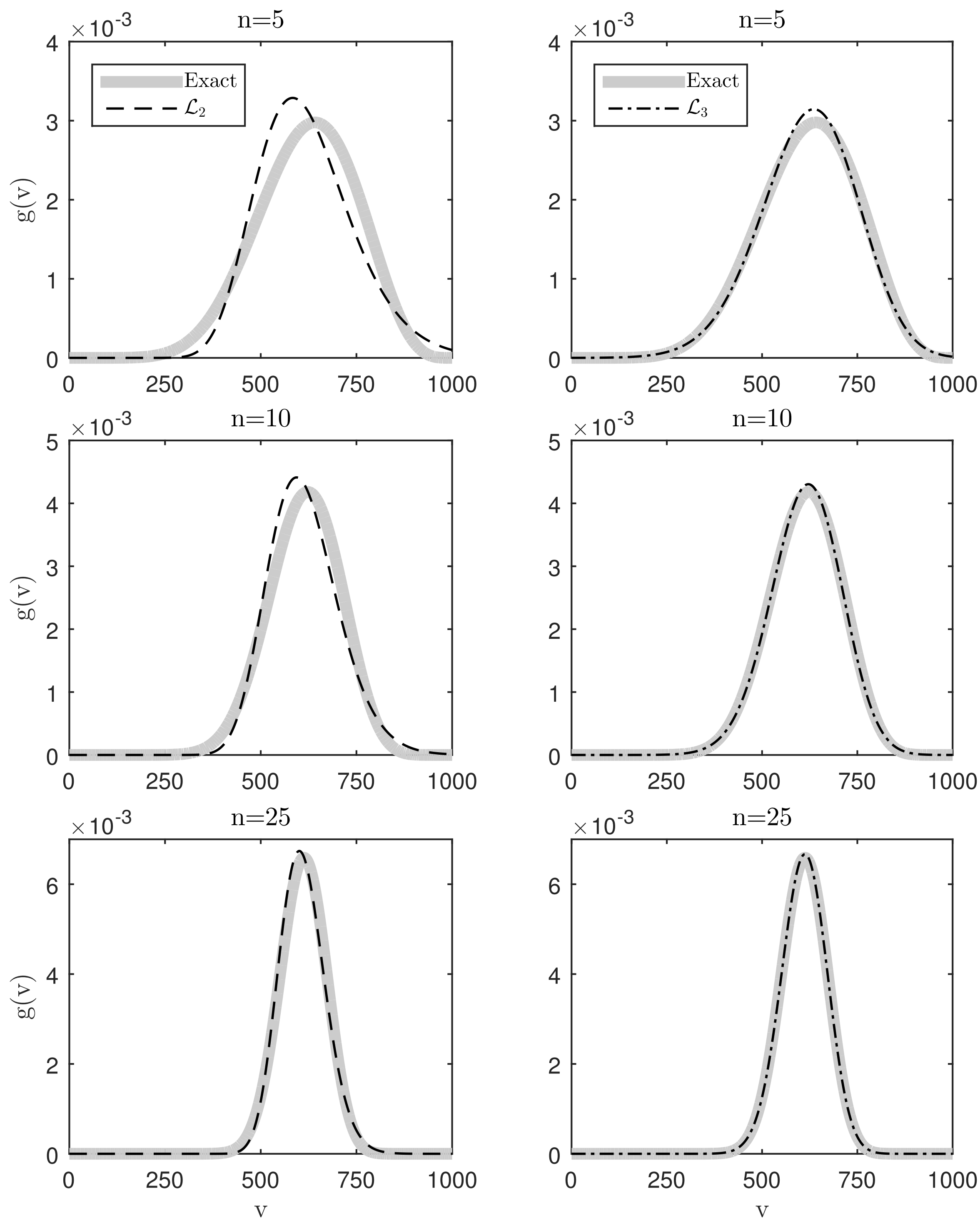
$$\alpha = 0.5\left(\ln\left(\frac{\sigma^2}{e^{\beta^2}-1}\right)-\beta^2\right), \tag{28}$$

$$\kappa = \mu-\delta e^{\alpha+0.5\beta^2}, \tag{29}$$

where μ, σ^2, and γ are the mean, variance, and skewness of the NPV distribution.

In order to illustrate the accuracy of the lognormal approximations, we revisit the last example of § 3. Figure 2 shows the exact and the approximate PDF of the NPV distribution for various values of n. Table 1 reports the mean, variance, skewness, kurtosis, and Kolmogorov-Smirnov test statistic. We observe that the $\mathcal{L}_3$ approximation is almost always very accurate, whereas the $\mathcal{L}_2$ approximation has more or less the same accuracy as the $\mathcal{L}_N$ approximation. This latter observation is no surprise. If n is small, neither Lemma 3 nor Theorem 3 hold, and the approximations fail to achieve a good accuracy. In addition, the $\mathcal{L}_2$ and $\mathcal{L}_N$ approximations only take into account the first two moments. As a result, they are always dominated by the $\mathcal{L}_3$ approximation.

Figure 2: PDF of the exact NPV, the $\mathcal{L}_2$, and the $\mathcal{L}_3$ approximation for various number of stages



5 NPV of a project with multiple stages and intermediate cash flows

In this section, we consider a project with multiple stages $w : w \in \mathbf{N} = \{1, 2, \ldots, n\}$, and assume that a cash flow c_w is incurred at the start of stage w. A payoff p is obtained upon completion of the project. For notational convenience, we let $c_{n+1} \equiv p$. Let $\mathbf{c} = \{c_1, c_2, \ldots, c_n, c_{n+1}\}$ denote the set of cash flows that are incurred during the lifetime of the project. In addition, define V_w, the random variable that represents the NPV of cash flow c_w, and let $V_{\mathbf{c}} = \sum_{w=1}^{n+1} V_w$ denote the random variable that captures the NPV of the project. Because the NPV of a cash flow c_x depends on the NPV of an earlier cash flow c_w, V_x depends on V_w for all $x, w : 1 \leq w < x \leq n+1$. Hence, $V_{\mathbf{c}}$ is the sum of a number of dependent random variables whose distribution converges to the lognormal distribution if their associated cash flow is not incurred during the early stages of the project.

Determining the distribution of $V_{\mathbf{c}}$ is closely related to finding the distribution of the lognormal sum (i.e., the sum of a number of random variables that follow a lognormal distribution). Even though the lognormal sum has received considerable attention in the literature (see Yan et al. (2016) for a brief overview), only few exact results are available. In what follows, we first develop exact, closed-form expressions for the moments of the distribution of $V_{\mathbf{c}}$. We then use the lognormal approximation developed in § 4 to approximate the NPV distribution, and assess its accuracy by means of an example. Next, we show that $V_{\mathbf{c}}$ is normally distributed if the number of cash flows is sufficiently large, and propose a new approximation based on the normal distribution. Again, we assess the accuracy of this approximation by means of an example.

Theorem 4. *Consider a project with multiple stages $w : w \in \mathbf{N}$, and let $\mathbf{c} = \{c_1, c_2, \ldots, c_n, c_{n+1}\}$ denote the set of cash flows that are incurred at the start of each stage (where $c_{n+1} \equiv p$ is the payoff that is obtained upon project completion). In addition, V_w denotes the random variable that represents the NPV of cash flow c_w, and $V_{\mathbf{c}} = \sum_{w=1}^{n+1} V_w$ is the random variable that captures the*

NPV of the project. The moments of the distribution of $V_{\mathbf{c}}$ are:

$$\begin{aligned}
\mu_{\mathbf{c}} &= \sum_{w=1}^{n+1} \mu_w, && (30)\\
\sigma_{\mathbf{c}}^2 &= \mathbf{e}\mathbf{\Sigma}_{\mathbf{c}}\mathbf{e}, && (31)\\
\gamma_{\mathbf{c}} &= (\mathbf{e}\mathbf{\Gamma}_{\mathbf{c}}\mathbf{e})\,\sigma_{\mathbf{c}}^{-3}, && (32)\\
\theta_{\mathbf{c}} &= (\mathbf{e}\mathbf{\Theta}_{\mathbf{c}}\mathbf{e})\,\sigma_{\mathbf{c}}^{-4}, && (33)
\end{aligned}$$

where $\mathbf{e}$ is a vector of ones, and $\mathbf{\Sigma}_{\mathbf{c}}$, $\mathbf{\Gamma}_{\mathbf{c}}$, and $\mathbf{\Theta}_{\mathbf{c}}$ are the central covariance, coskewness, and cokurtosis matrices respectively. $\mathbf{\Sigma}_{\mathbf{c}}$, $\mathbf{\Gamma}_{\mathbf{c}}$, and $\mathbf{\Theta}_{\mathbf{c}}$ capture the covariance, coskewness, and cokurtosis of the NPV of the cash flows in $\mathbf{c}$. Table 3 provides a summary of the closed-form expressions that allow to calculate the entries of these cross-moment matrices.

In order to illustrate Theorem 4, we use an example project with 3 stages. In the example, cash outflows are incurred at the start of the project, and at the start of the third stage. Cash inflows, on the other hand, are received at the start of the second stage, and upon completion of the project. Each stage w has a duration that follows a gamma distribution with shape and scale parameters k_w and τ_w respectively. We assume a discount rate $r = 0.05$. The data of the example project are summarized in Table 4. Figure 3 shows the $\mathcal{L}_2$ and $\mathcal{L}_3$ approximations, as well as the simulated PDF of the project NPV (note that we have to resort to simulation as it is no longer an easy task to determine the exact NPV distribution). It is clear that, in this example, the $\mathcal{L}_2$ approximation performs very poorly. The $\mathcal{L}_3$ approximation, however, is once more very accurate. Table 5 reports on the moments of the NPV distribution, the probability to have a negative project NPV, and the Kolmogorov–Smirnov test statistic. The exact moments have been obtained using Theorem 4. We observe that the simulation (with 1 billion replications) almost perfectly matches the exact moments, which supports the claim that the simulated PDF is close to the true PDF of the project NPV. As was also shown by Figure 3, the $\mathcal{L}_3$ approximation yields excellent accuracy. If, however, cross moments are ignored (i.e., if we assume that the NPVs of the cash flows are independent), the accuracy is abysmal. This is also reflected in the probability to have a negative project NPV. Only the $\mathcal{L}_3$ approximation is able to provide an accurate estimate.

Theorem 5. *Consider a project with multiple stages $w : w \in \mathbf{N} = \{1, 2, \ldots, n\}$ that are executed*

Mean μ
$\mu_w = c_w a_1$

covariance matrix $\mathbf{\Sigma_c}$
$\mathbf{\Sigma_c}(w,w) = \sigma_w^2 = c_w^2(a_2 - a^2)$
$\mathbf{\Sigma_c}(w,x) = c_w c_x b_1 \left(a_2 - a^2\right) = c_w^{-1} c_x b_1 \mathbf{\Sigma_c}(w,w)$

Central coskewness matrix $\mathbf{\Gamma_c}$
$\mathbf{\Gamma_c}(w,w,w) = \gamma_w \sigma_w^3 = c_w^3 \left(a_3 - 3a_2 a_1 + 2a^3\right)$
$\mathbf{\Gamma_c}(w,w,x) = c_w^{-1} c_x b_1 \mathbf{\Gamma_c}(w,w,w)$
$\mathbf{\Gamma_c}(w,x,x) = c_w c_x^2 \left(a_3 b_2 - a_2 a_1 \left(2b^2 + b_2\right) + 2a^3 b^2\right)$
$\mathbf{\Gamma_c}(w,x,y) = c_x^{-1} c_y h_1 \mathbf{\Gamma_c}(w,x,x)$

Central cokurtosis matrix $\mathbf{\Theta_c}$
$\mathbf{\Theta_c}(w,w,w,w) = \theta_w \sigma_w^4 = c_w^4 \left(a_4 - 4a_3 a_1 + 6a_2 a^2 - 3a^4\right)$
$\mathbf{\Theta_c}(w,w,w,x) = c_w^{-1} c_x b_1 \mathbf{\Theta_c}(w,w,w,w)$
$\mathbf{\Theta_c}(w,w,x,x) = c_w^2 c_x^2 \left(a_4 b_2 - 2a_3 a_1 \left(b_2 + b^2\right) + a_2 a^2 \left(b_2 + 5b^2\right) - 3a^4 b^2\right)$
$\mathbf{\Theta_c}(w,x,x,x) = c_w c_x^3 \left(a_4 b_3 - a_3 a_1 \left(b_3 + 3b_2 b_1\right) + 3a_2 a^2 \left(b_2 b_1 + b^3\right) - 3a^4 b^3\right)$
$\mathbf{\Theta_c}(w,w,x,y) = c_x^{-1} c_y h_1 \mathbf{\Theta_c}(w,w,x,x)$
$\mathbf{\Theta_c}(w,x,x,y) = c_x^{-1} c_y h_1 \mathbf{\Theta_c}(w,x,x,x)$
$\mathbf{\Theta_c}(w,x,y,y) = c_w c_x c_y^2 \left(\left(a_4 - a_3 a_1\right) b_3 h_2 - \left(h_2 + 2h^2\right) \left(\left(a_3 a_1 - a_2 a^2\right) b_2 b_1\right) + \left(a_2 a^2 - a^4\right) 3b^3 h^2\right)$
$\mathbf{\Theta_c}(w,x,y,z) = c_y^{-1} c_z o_1(r) \mathbf{\Theta_c}(w,x,y,y)$

$a_i = \phi_{1,w-1}(ir)$	$b_i = \phi_{w,x-1}(ir)$	$h_i = \phi_{x,y-1}(ir)$	$o_i = \phi_{y,z-1}(ir)$
$a^i = \phi^i_{1,w-1}(r)$	$b^i = \phi^i_{w,x-1}(r)$	$h^i = \phi^i_{x,y-1}(r)$	

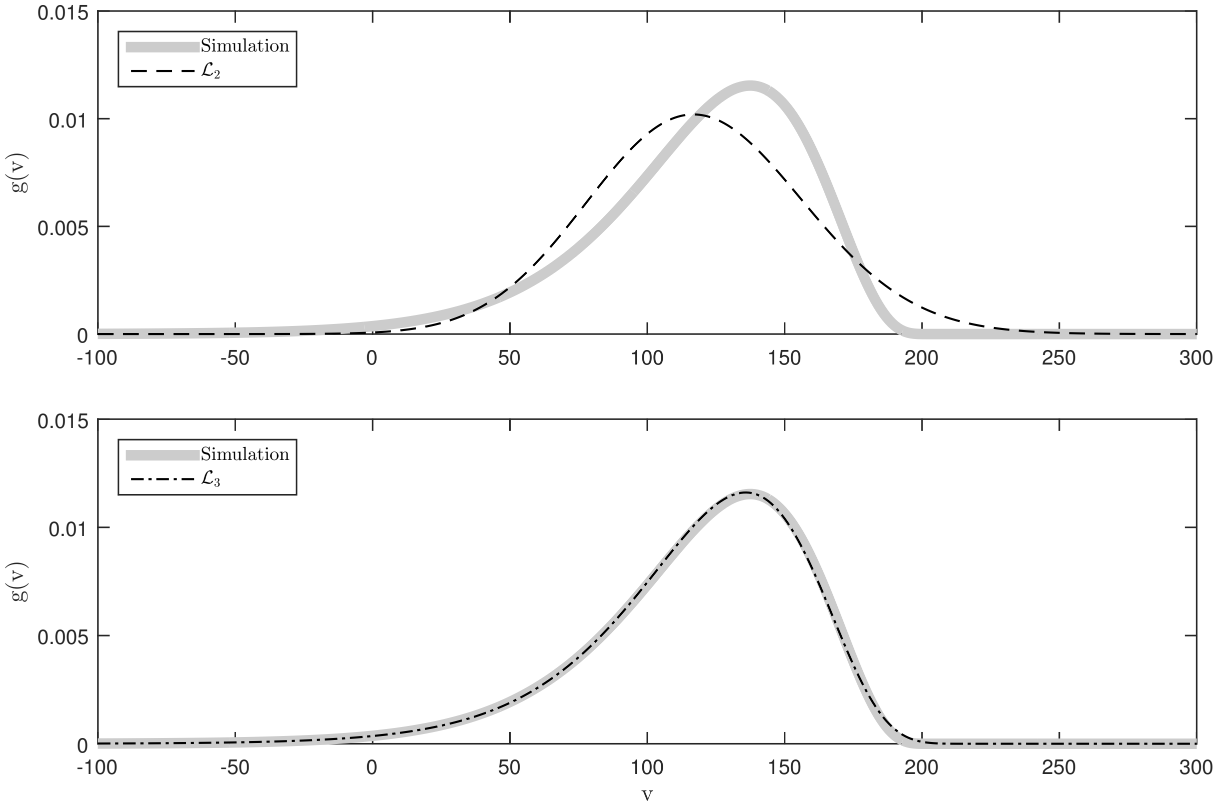
Table 3: Summary of closed-form expressions that allow to calculate the moments of the NPV distribution of a project

w	c_w	k_w	τ_w	d_w	s_w^2
1	-300	1.5	1.0	1.5	1.5
2	250	2.5	1.0	2.5	2.5
3	-750	0.5	1.0	0.5	0.5

p	1000
r	0.05

Table 4: Data of the example project with three stages

Figure 3: PDF of the simulated NPV, the $\mathcal{L}_2$, and the $\mathcal{L}_3$ approximation for a project with intermediate cash flows



	Exact	Simulation	$\mathcal{L}_2$	$\mathcal{L}_3$	$\mathcal{L}_3$ without cross moments
μ	168.21	168.21	168.21	168.21	168.21
σ^2	1,533	1,533	1,533	1,533	10,276
γ	-1.035	-1.035	0.1006	-1.035	-2.620
θ	4.7421	4.7420	3.0180	4.9631	17.269
$G(0)$	NA	0.0105	0.0008	0.0105	0.1018
K–S	NA	NA	0.0734	0.0055	0.1018

Table 5: Accuracy of the $\mathcal{L}_2$ and $\mathcal{L}_3$ approximations of the NPV distribution of a project with intermediate cash flows

in sequence. At the start of each stage $w : w \in \mathbf{N}$, a cash flow c_w is incurred, and a payoff $p \equiv c_{n+1}$ is obtained upon completion of the project. Let V_w denote the random variable that represents the NPV of cash flow c_w, and let $V_{\mathbf{c}} = \sum_{w=1}^{n+1} V_w$ denote the random variable that captures the NPV of the project. If $r > 0$, and if $s_w^2 > 0$ for all $w \in \mathbf{N}$, the project NPV converges to a normal distribution, with mean $\mu_{\mathbf{c}}$ and variance $\sigma_{\mathbf{c}}^2$, as the number of stages increases.

Note that Theorem 5 also applies in a more general context where stages are not necessarily executed in sequence. In fact, Theorem 5 holds as long as a sufficient number of cash flows are incurred during the lifetime of a project.

We use an example to illustrate Theorem 5. The example project has n stages with gamma-distributed durations with shape and scale parameter k_i and τ_i respectively. Cash outflows are incurred at the start of uneven stages. Cash inflows, on the other hand, are obtained at the start of even stages, and upon completion of the project. The discount rate r equals $0.1n^{-1}$. Table 6 summarizes the data of the example project. Figure 4 shows the simulated and the approximate PDF of the distribution of the project NPV. Next to the lognormal $\mathcal{L}_3$ approximation, we now also include a normal approximation that has mean $\mu_{\mathbf{c}}$ and variance $\sigma_{\mathbf{c}}^2$, and that is denoted by $\mathcal{N}$. We observe that, as n increases, the project NPV converges to a normal distribution, and the accuracy of the $\mathcal{N}$ approximation improves. Even so, the $\mathcal{L}_3$ approximation still performs better due to the extra moment matched. These findings are confirmed by Table 7 that reports on the moments of the NPV distribution, and on the Kolmogorov–Smirnov test statistic. For reference, we have also included the CPU time required to run the simulation (with 1 billion replications) and to calculate the moments using the closed-form expressions provided in Table 3. Both the simulation as well as the exact approach were implemented in Visual Studio C++. Although the simulation model yields good accuracy (also for a lower number of replications), it can hardly compete with an exact, closed-form approach that requires less than a second of CPU time when 100 stages are considered. In addition, most of the computation time is spent on calculating the cokurtosis matrix. Our approach, however, only requires that the first three moments are specified (i.e., there is no need to determine the kurtosis). If only three moments are calculated, the required CPU time drops to 0.046 seconds (for $n = 100$). Even if, for very large n, the computation of the coskewness matrix becomes too time consuming, it suffices to calculate only the first two moments as the $\mathcal{N}$

$$c_w = \begin{cases} 250 & \text{if } w \text{ is even} \\ -250 & \text{if } w \text{ is uneven} \end{cases}$$

$$k_w = \begin{cases} 0.5 & \text{if } w \in \{1, 6, 11, \ldots\} \\ 1.0 & \text{if } w \in \{2, 7, 12, \ldots\} \\ 1.5 & \text{if } w \in \{3, 8, 13, \ldots\} \\ 2.0 & \text{if } w \in \{4, 9, 14, \ldots\} \\ 2.5 & \text{if } w \in \{5, 10, 15, \ldots\} \end{cases}$$

$$\tau_w = \begin{cases} 2.0 & \text{if } w \text{ is even} \\ 1.0 & \text{if } w \text{ is uneven} \end{cases}$$

$$p = 1000$$
$$r = 0.1n^{-1}$$

Table 6: Data of the example project with n stages and intermediate cash flows

	$n = 10$			$n = 30$			$n = 100$		
	Sim	$\mathcal{N}$	$\mathcal{L}_3$	Sim	$\mathcal{N}$	$\mathcal{L}_3$	Sim	$\mathcal{N}$	$\mathcal{L}_3$
μ	783.04	783.04	783.04	782.16	782.16	782.16	781.86	781.86	781.86
σ^2	2,584	2,584	2,584	875	875	875	264	264	264
γ	-0.361	0.0	-0.361	-0.211	0.0	-0.211	-0.117	0.0	-0.116
θ	3.1159	3.0	3.1162	3.0415	3.0	3.0402	3.0769	3.0	3.0122
K–S	NA	0.0297	0.0032	NA	0.0155	0.0001	NA	0.0080	0.0003
CPU (s)	2,250	0.000		11,279	0.015		90,955	0.967	

Table 7: Accuracy of the $\mathcal{N}$ and $\mathcal{L}_3$ approximations for the NPV of a project with intermediate cash flows and n stages

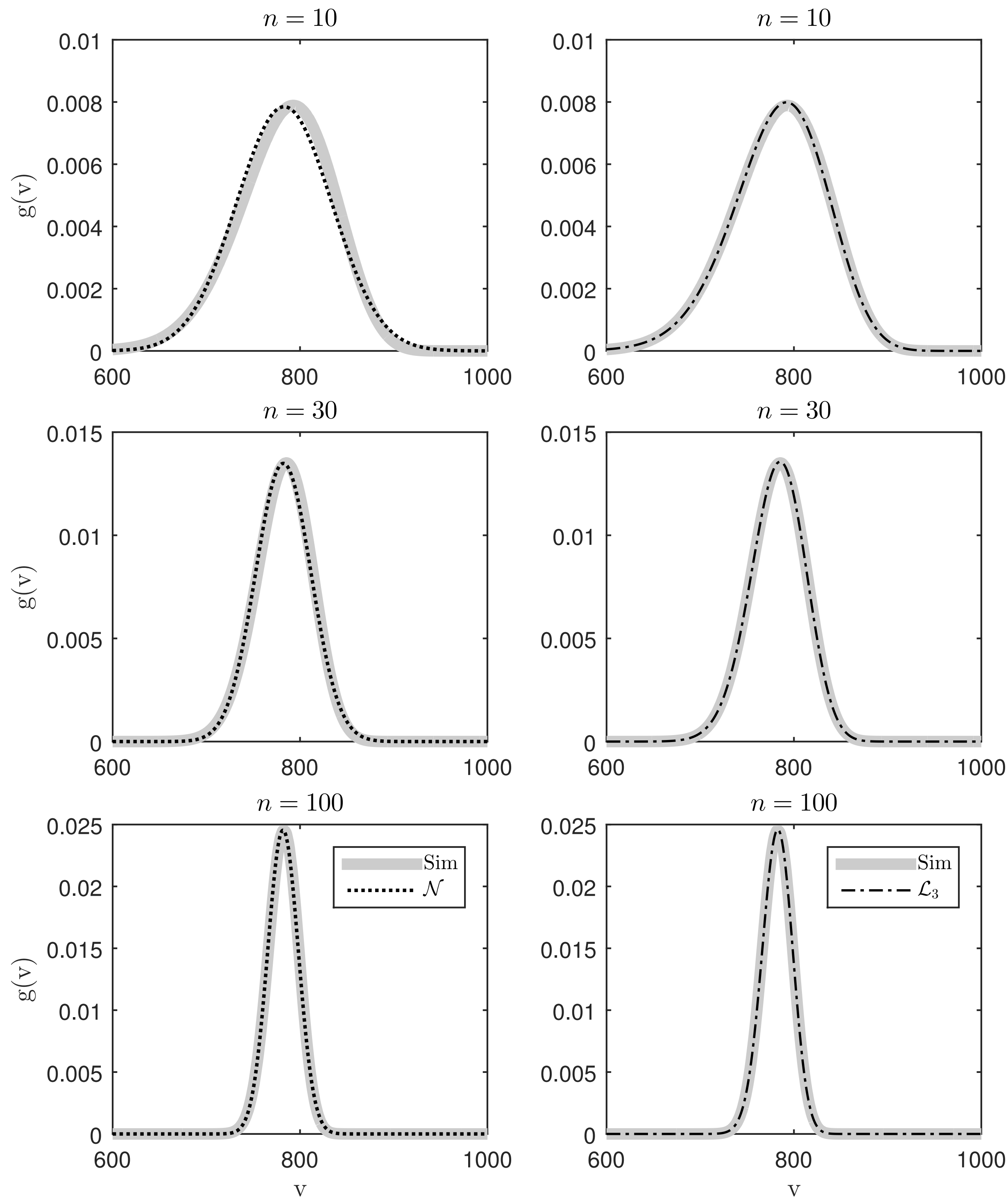
approximation becomes more accurate as n increases (i.e., for large n, it is no longer necessary to calculate the coskewness matrix).

6 Model extensions

In this section, we discuss two model extensions. A first extension allows stages (and hence projects) to fail. Stage/project failure is common in the literature on R&D projects (see e.g., Creemers et al. (2015)), and can easily be incorporated in our approach. We need only to redefine factor $\phi_w(r)$:

$$\phi_w(r) = j_w \phi_w^*(r), \tag{34}$$

Figure 4: PDF of the simulated NPV, the $\mathcal{N}$, and the $\mathcal{L}_3$ approximation for various number of stages



where j_w is the probability of success of stage w, and $\phi^*_w(r)$ is the factor given by (4) (i.e., the factor that does not take into account stage/project failure).

A second model extension allows for different discount rates to be applied during different stages of the project. This extension requires a redefinition of factor $\phi_{w,x}(r)$:

$$\phi_{w,x}(\mathbf{r}) = \prod_{y=w}^{x} \phi_y(r_y), \tag{35}$$

where r_y is the discount rate that applies for stage y, and $\mathbf{r} = \{w, w+1, \ldots, x\}$ is the vector of discount rates that apply to stages $y : w \leq y \leq x$.

7 Conclusions

In this article, we considered projects with multiple stages $w : w \in \mathbf{N} = \{1, 2, \ldots, n\}$ that are executed in sequence. Each stage $w : w \in \mathbf{N}$ has a random duration T equipped with probability function $f(t)$. A cash flow c_w (positive or negative) may be incurred upon the start of stage w, and a payoff p is obtained at the completion of the project. We use continuous compounding and a discount rate r to determine the NPV of a project.

Our main contributions can be summarized as follows: (1) we obtain exact, closed-form expressions for the moments of the NPV of a project, (2) we develop a highly accurate closed-form approximation of the distribution of the project NPV, (3) we show that the NPV of a single cash flow converges to a lognormal distribution if the cash flow is not incurred during the early stages of the project, and (4) we show that the NPV of a project converges to a normal distribution if a sufficient number of cash flows are incurred during the lifetime of the project. Contributions (3) and (4) are valid in a more general setting (i.e., in a setting where stages are not necessarily executed in sequence).

Our work can directly be applied in the fields of project selection, project portfolio management, and project valuation. In these fields, a project is often seen as a sequence of stages with intermediate cash flows (including a payoff that is obtained upon the successful completion of the project). Project selection/investment decisions can be made based on the eNPV and the risk of a project. The risk of a project/an investment is often modeled using the variance of the NPV.

Other measures of risk include the skewness and/or kurtosis of the NPV, and the probability to have a negative NPV. Until now, Monte Carlo simulation was the only tool available to obtain estimates for these measures. Our work, however, renders the use of Monte Carlo simulation obsolete, and allows to obtain a highly accurate approximation of the NPV distribution, and an exact characterization of its moments.

In the (more operational) field of project scheduling, our work is related to CPM/PERT in the sense that we also focus on a single sequence of stages, and also use normal (lognormal) approximations. As a result, the limitations of our work are similar to those of CPM/PERT. Future research should focus on generalizing the problem setting by allowing stages to be executed in parallel rather than only in sequence. Methods that have been used to generalize CPM/PERT may also be applied here (e.g., network transformations/reductions and bounding procedures). In addition, we have assumed that the sequence of stages is known in advance (i.e., no scheduling decisions are made). In reality, however, the sequence of stages is not always predetermined. Therefore, the optimal sequence of stages is another direction for future research. Note that this research topic is related to the literature that deals with the optimal sequence of tests (where tests are the stages of the project, and the probability of technical success of a stage w is given by factor $\phi_w(r)$).

Appendix. Proofs

PROOF OF LEMMA 1. The proof follows from the definition of the MGF:

$$\begin{aligned} M_T(u) &= \int_0^\infty f(t)e^{ut}\mathrm{d}t, \quad \text{if } T \text{ is continuous,} \\ &= \sum_{t=0}^{\infty} f(t)e^{ut}, \quad \text{if } T \text{ is discrete.} \end{aligned} \tag{36}$$

□

PROOF OF THEOREM 1. Let V denote the random variable that represents the NPV of a cash flow

c that is incurred at time T. The MGF of V is:

$$M_V(u) = \sum_{i=0}^{\infty} \frac{u^i m_i}{i!}, \tag{37}$$

where m_i is the ith raw moment of the NPV distribution:

$$\begin{aligned} m_i &= \int_0^{\infty} f(t)(e^{-rt})^i \mathrm{d}t = \phi(ir), \quad \text{if } T \text{ is continuous,} \\ &= \sum_{t=0}^{\infty} f(t)(e^{-rt})^i = \phi(ir), \quad \text{if } T \text{ is discrete.} \end{aligned} \tag{38}$$

Using these raw moments, we can obtain the mean, variance, skewness, kurtosis, and even higher-order moments of the NPV of cash flow c. □

Proof of Theorem 2. If we solve (1) for t, we obtain:

$$t_v = \ln\left(\frac{c}{v}\right) r^{-1}. \tag{39}$$

where t_v is the time at which cash flow c needs to be incurred in order to obtain NPV v for a given discount rate r. As a result, $F(t_v)$ not only represents the probability to have a time $t \leq t_v$, but it also represents the probability to have an NPV $\geq v$. □

Proof of Lemma 2. Factor $\phi_{1,n}(r)$ can be obtained as follows:

$$\begin{aligned} \phi_{1,n}(r) &= \int_0^{\infty} \cdots \int_0^{\infty} f_1(t_1)e^{-rt_1} \cdots f_n(t_n)e^{-rt_n} \mathrm{d}t_1 \cdots \mathrm{d}t_n, \\ &= \phi_1(r) \int_0^{\infty} \cdots \int_0^{\infty} f_2(t_2)e^{-rt_2} \cdots f_n(t_n)e^{-rt_n} \mathrm{d}t_2 \cdots \mathrm{d}t_n, \\ \cdots \\ &= \prod_{w \in \mathbf{N}} \phi_w(r). \end{aligned} \tag{40}$$

In general, let $\phi_{w,x}(r)$ denote the factor for stages w to x, where $x \geq w$:

$$\phi_{w,x}(r) = \prod_{y=w}^{x} \phi_y(r). \tag{41}$$

$\square$

Proof of Lemma 3. The proof follows from the Central Limit Theorem (CLT). $\square$

Proof of Theorem 3. The proof is a direct application of Theorem 2 and Lemma 3. If n is sufficiently large, the duration of the project is normally distributed, and if $F(t)$ is a normal distribution function, $G(v)$ can be expressed as follows:

$$G(v) = \frac{1}{2} + \frac{1}{2}\,\mathrm{Erf}\left(\frac{\ln(v) - (\ln(p) - rd_{\mathbf{N}})}{\sqrt{2}rs_{\mathbf{N}}}\right). \tag{42}$$

When substituting $\ln(p) - rd_{\mathbf{N}}$ by α and $rs_{\mathbf{N}}$ by β, we get:

$$G(v) = \frac{1}{2} + \frac{1}{2}\,\mathrm{Erf}\left(\frac{\ln(v) - \alpha}{\sqrt{2}\beta}\right), \tag{43}$$

which is the CDF of the lognormal distribution with location parameter $\alpha = \ln(p) - rd_{\mathbf{N}}$ and scale parameter $\beta = rs_{\mathbf{N}}$. $\square$

Proof of Lemma 4. Consider a lognormal distribution with location and scale parameter α and β respectively. The mean and SCV of that distribution are given by:

$$\mu_{\mathcal{L}2} = e^{\alpha + 0.5\beta^2}, \tag{44}$$

$$\eta^2_{\mathcal{L}2} = e^{(2\alpha+\beta^2)}\left(e^{\beta^2} - 1\right)e^{-2(\alpha+0.5\beta^2)}. \tag{45}$$

The unique solution for β can easily be obtained by solving (45):

$$\beta = \sqrt{\ln(1 + \eta^2_{\mathcal{L}2})}. \tag{46}$$

Note that, as long as $\sigma^2_{\mathcal{L}2} > 0$, then $\eta^2_{\mathcal{L}2} > 0$, and therefore $\beta > 0$. Given β, the unique solution for α can easily be obtained by solving (44):

$$\alpha = \ln(\mu_{\mathcal{L}2}) - 0.5\beta^2. \tag{47}$$

We conclude that α and β are the unique solution to (44–45), and that they are well defined for all

$\mu_{\mathcal{L}2}, \sigma^2_{\mathcal{L}2} > 0$. □

Proof of Lemma 5. First, we obtain the unique solution for β from (23). We have:

$$\gamma_{\mathcal{L}3} = \delta(2 + e^{\beta^2})\sqrt{e^{\beta^2} - 1}, \tag{48}$$

$$\gamma^2_{\mathcal{L}3} = e^{3\beta^2} + 3e^{2\beta^2} - 4. \tag{49}$$

If we let $q = e^{\beta^2}$ and $l = 4 + \gamma^2_{\mathcal{L}3}$, we obtain the following cubic equation:

$$q^3 + 3q^2 - l = 0. \tag{50}$$

The discriminant of $q^3 + 3q^2 - l$ is $\Delta = -27(l-4)l$, and is always negative if $\gamma_{\mathcal{L}3} \neq 0$. For cubic equations, if $\Delta < 0$, the equation has one unique real root and two non-real complex conjugate roots. The unique real root of $q^3 + q^2 - l$ is:

$$q = \frac{2^{1/3}}{\left(-2 + l + \sqrt{-4l + l^2}\right)^{1/3}} + \frac{\left(-2 + l + \sqrt{-4l + l^2}\right)^{1/3}}{2^{1/3}} - 1. \tag{51}$$

After substituting $q = e^{\beta^2}$ and $l = 4 + S^2_{\mathcal{L}3}$, we obtain the unique, real solution for β:

$$\beta = \sqrt{\ln\left(\frac{2^{1/3}}{\left(2 + \gamma^2_{\mathcal{L}3} + \sqrt{4\gamma^2_{\mathcal{L}3} + \gamma^4_{\mathcal{L}3}}\right)^{1/3}} + \frac{\left(2 + \gamma^2_{\mathcal{L}3} + \sqrt{4\gamma^2_{\mathcal{L}3} + \gamma^4_{\mathcal{L}3}}\right)^{1/3}}{2^{1/3}} - 1\right)}. \tag{52}$$

Given β, the unique solution for α can easily be obtained by solving (22):

$$\alpha = 0.5\left(\ln\left(\frac{\sigma^2_{\mathcal{L}3}}{e^{\beta^2} - 1}\right) - \beta^2\right). \tag{53}$$

Given α and β, the unique solution for κ can easily be obtained by solving (21):

$$\kappa = \mu_{\mathcal{L}3} - \delta e^{\alpha + 0.5\beta^2}. \tag{54}$$

We conclude that α, β, and κ are the unique solution to (21–23), and that they are well defined for all $\mu_{\mathcal{L}3}, \sigma^2_{\mathcal{L}3} > 0$, and for $\gamma_{\mathcal{L}3} \neq 0$. $\square$

Proof of Theorem 4. The covariance between the NPV of cash flow c_x and the NPV of an earlier cash flow c_w is given by:

$$\mathbf{\Sigma_c}(w,x) = \\ \int_0^\infty \!\!...\!\! \int_0^\infty \prod_{y=1}^{x-1} f_y(t_y) \left(c_w e^{-r\left(\sum_{y=1}^{w-1} t_y\right)} - \mu_w \right) \left(c_x e^{-r\left(\sum_{y=1}^{x-1} t_y\right)} - \mu_x \right) \mathrm{d}t_1 ... \mathrm{d}t_{x-1}. \tag{55}$$

Which can be rewritten as a sum of 4 parts:

$$\begin{aligned} &\mathbf{\Sigma_c}(w,x) = \\ &\int_0^\infty \cdots \int_0^\infty \prod_{y=1}^{x-1} f_y(t_y) \left(c_w e^{-r\left(\sum_{y=1}^{w-1} t_y\right)} c_x e^{-r\left(\sum_{y=1}^{x-1} t_y\right)} \right) \mathrm{d}t_1 \cdots \mathrm{d}t_{x-1} \\ &- \int_0^\infty \cdots \int_0^\infty \prod_{y=1}^{x-1} f_y(t_y) \left(\mu_x c_w e^{-r\left(\sum_{y=1}^{w-1} t_y\right)} \right) \mathrm{d}t_1 \cdots \mathrm{d}t_{x-1} \\ &- \int_0^\infty \cdots \int_0^\infty \prod_{y=1}^{x-1} f_y(t_y) \left(\mu_w c_x e^{-r\left(\sum_{y=1}^{x-1} t_y\right)} \right) \mathrm{d}t_1 \cdots \mathrm{d}t_{x-1} \\ &+ \int_0^\infty \cdots \int_0^\infty \prod_{y=1}^{x-1} f_y(t_y) \left(\mu_w \mu_x \right) \mathrm{d}t_1 \cdots \mathrm{d}t_{x-1}. \end{aligned} \tag{56}$$

After application of Lemma 2, we get:

$$\begin{aligned} &\mathbf{\Sigma_c}(w,x) = \\ &c_w c_x \phi_{1,w-1}(2r)\phi_{w,x-1}(r) \\ &-\mu_x c_w \phi_{1,w-1}(r) \\ &-\mu_w c_x \phi_{1,w-1}(r)\phi_{w,x-1}(r) \\ &+\mu_w \mu_x. \end{aligned} \tag{57}$$

From Theorem 1, we have that $\mu_w = c_w\phi_{1,w-1}(r)$ and $\mu_x = c_x\phi_{1,x-1}(r)$, and therefore:

$$
\begin{aligned}
&\mathbf{\Sigma_c}(w,x) = \\
&c_w c_x \phi_{1,w-1}(2r)\phi_{w,x-1}(r) \\
&-c_w c_x \phi_{1,w-1}(r)\phi_{1,w-1}(r)\phi_{w,x-1}(r) \\
&-c_w c_x \phi_{1,w-1}(r)\phi_{1,w-1}(r)\phi_{w,x-1}(r) \\
&+c_w c_x \phi_{1,w-1}(r)\phi_{1,w-1}(r)\phi_{w,x-1}(r).
\end{aligned} \tag{58}
$$

Which, finally, can be simplified to:

$$
\mathbf{\Sigma_c}(w,x) = c_w c_x \phi_{w,x-1}(r)\left(\phi_{1,w-1}(2r) - \phi^2_{1,w-1}(r)\right). \tag{59}
$$

The same approach can be used to determine the coskewness, the cokurtosis, and even the higher-order cross moments of the NPV of the cash flows that are incurred during the lifetime of the project. □

Proof of Theorem 5. Let $(V) = \{V_1, V_2, \ldots, V_n, V_{n+1}\}$ denote the non-stationary sequence of dependent random variables $V_w : 1 \leq w \leq n+1$. For such a sequence, Bradley and Tone (2015) have shown that a CLT holds if:

- the sequence is strongly mixing,
- the sequence has a maximum correlation that is strictly smaller than 1 for some V_w and V_{w+1} in (V),
- the Lindeberg condition holds.

Several mixing conditions have been defined in the literature (for an overview, see Bradley (2005)). In this proof, we will show that sequence (V) is ρ-mixing (which automatically implies that (V) is strongly mixing). A sequence is said to be ρ-mixing if the maximum correlation between two random variables $V_w, V_x \in (V)$ tends to zero for some w and x that are "far apart". We use (59) to obtain the expression for the correlation between two random variables V_w and V_x:

$$
\begin{aligned}
\mathrm{Corr}(w,x) &= \frac{\phi_{w,x-1}(r)\left(\phi_{1,w-1}(2r)-\phi^2_{1,w-1}(r)\right)}{\sqrt{\phi_{1,w-1}(2r)-\phi^2_{1,w-1}(r)}\sqrt{\phi_{1,x-1}(2r)-\phi^2_{1,x-1}(r)}}, \\
&= \phi_{w,x-1}(r)\sqrt{\frac{\phi_{1,w-1}(2r)-\phi^2_{1,w-1}(r)}{\phi_{1,x-1}(2r)-\phi^2_{1,x-1}(r)}}. \qquad (60)
\end{aligned}
$$

It is easy to verify that $\mathrm{Corr}(w,x) \to 0$ if $\phi_{w,x-1}(r) \to 0$, or if $\phi_{1,w-1}(2r) = \phi^2_{1,w-1}(r)$. If $c_w > 0$, and if at least one stage $z : 1 \le z < w$ has $s_z^2 > 0$, then $\sigma_w > 0$, and it follows from (6) that $\phi_{1,w-1}(2r) > \phi^2_{1,w-1}(r)$. Therefore, we say that $\mathrm{Corr}(w,x) \to 0$ if and only if $\phi_{w,x-1}(r) \to 0$. From Lemma 2 we know that $\phi_{w,x-1}(r) = \prod_{y=w}^{x-1} \phi_y(r)$. In addition, if $r > 0$, and if $s_y^2 > 0$, then $\phi_y(r) < 1$, and $\phi_{w,x-1}(r) \to 0$ if $s_y^2 > 0$ for sufficient $y : w \le y < x$, and for $x - w \to \infty$. In other words, if $r > 0$, and if $s_y^2 > 0$ for sufficient $y \in \mathbf{N}$, then sequence (V) is ρ-mixing as $n \to \infty$.

In order to show that sequence (V) satisfies the second condition, we observe the correlation between random variables V_w and V_{w+1}:

$$
\begin{aligned}
\mathrm{Corr}(w,w+1) &= \phi_w(r)\sqrt{\frac{\phi_{1,w-1}(2r)-\phi^2_{1,w-1}(r)}{\phi_{1,w}(2r)-\phi^2_{1,w}(r)}}, \\
&= \frac{\phi_{1,w-1}(2r)\phi^2_w(r)-\phi^2_{1,w}(r)}{\phi_{1,w-1}(2r)\phi_w(2r)-\phi^2_{1,w}(r)}. \qquad (61)
\end{aligned}
$$

A perfect correlation is achieved if $\phi_{1,w-1}(2r) \to 0$, or if $\phi_w(2r) = \phi^2_w(r)$. If $s_w^2 > 0$, then $\phi_w(2r) > \phi^2_w(r)$, and as a result, $\mathrm{Corr}(w,w+1) \to 1$ if and only if $\phi_{1,w-1}(2r) \to 0$. If $w \to \infty$, $\phi_{1,w-1}(2r) \to 0$, however, because $\phi_{1,w-1}(2r) > \phi^2_{1,w-1}(r)$, $\phi^2_{1,w-1}(r)$ goes to zero even faster. In addition, $\phi_w(2r) > \phi^2_w(r)$, and therefore, the maximum correlation between random variables V_w and V_{w+1} is always strictly smaller than 1.

To complete the proof, we still need to show that the Lindeberg condition holds. Instead of verifying the Lindeberg condition itself, we show that sequence (V) satisfies the more strict Lyapunov condition. The Lyapunov condition requires that all random variables $V_w \in (V)$ have finite mean, variance, and at least one finite higher-order moment (see e.g., Greene 2003). In our case, the ith moment of V_x is finite if $\phi_{1,x-1}(ir)$ is finite; if the MGF about $-ir$ is defined for all duration distributions $f_w(t) : 1 \le w < x$ (see also Lemma 1). In general, the MGF is defined for most duration distributions, and for most values of r. Therefore, we conclude that, in general, the

Lyapunov condition (and hence the Lindeberg condition) holds. □

References

Aitchison J, Brown JAC (1957) *The Lognormal Distribution, with Special Reference to Its Use in Economics.* (Cambridge University Press, New York).

Boute R, Lambrecht M, Van Houdt B (2007) Performance evaluation of a production/inventory system with periodic review and endogenous lead times. *Naval Res. Logist.* 54(4):462–473.

Bradley RC (2005) Basic properties of strong mixing conditions. A survey and some open questions. *Probab. Surveys.* 2(2005):107–144.

Bradley RC, Cristina T (2015) A central limit theorem for non-stationary strongly mixing random fields. *J. Theor. Probab.* 2015:1–20.

Creemers S, Leus R, Lambrecht M (2010) Scheduling Markovian PERT networks to maximize the net present value. *Oper. Res. Lett.* 38(1):51-56.

Creemers S, Demeulemeester E, Van de Vonder S (2014) A new approach for quantitative risk analysis. *Ann. Oper. Res.* 213(1):27–65.

Creemers S, Leus R, De Reyck B (2015) Project planning with alternative technologies in uncertain environments. *Eur. J. Oper. Res.* 242(2):465–476.

Demeulemeester E, Herroelen W (2002) *Project Scheduling: A Research Handbook.* (Kluwer Academic Publishers, Dordrecht).

De Reyck B, Degraeve Z, Vandenborre R (2008) Project options valuation with net present value and decision tree analysis. *Eur. J. Oper. Res.* 184(1):341–355.

Girotra K, Terwiesch C, Ulrich KT (2007) Valuing R&D projects in a portfolio: Evidence from the pharmaceutical industry. *Management Sci.* 53(9):1452–1466.

Hagstrom JN (1988) Computational complexity of PERT problems. *Networks.* 18(2):139-147.

Huchzermeier A, Loch CH (2001) Project management under risk: Using the real options approach to evaluate flexibility in R&D. *Management Sci.* 47(1):85–101.

Kelley JE Jr, Walker MR (1959) Critical-path planning and scheduling. *Proc. Eastern Joint Computer Conf., December 13, Boston,* 160-173.

Kwak YH, Ingall L (2007) Exploring Monte Carlo simulation applications for project management. *Risk Management.* 9(1):44–57.

Malcolm DG, Rosenboom JH, Clark CE, Fazar W (1959) Application of a technique for research and development program evaluation. *Oper. Res.* 7(5):646–669.

Moder JJ, Phillips CR (1970) *Project Management with CPM and PERT*, 2nd ed. (Van Nostrand Reinhold Company, New York).

Osogami T (2005) *Analysis of multiserver systems via dimensionality reduction of Markov chains*, PhD thesis. (Carnegie Mellon University, Pittsburgh).

Osogami T, Harchol-Balter M (2006) Closed form solutions for mapping general distributions to quasi-minimal PH distributions. *Performance Evaluation.* 63(6):524-552.

Savvides S (1994) Risk analysis in investment appraisal. *Project Appraisal.* 9(1):3–18.

Santiago LP, Vakili P (2005) On the value of flexibility in R&D projects. *Management Sci.* 51(8):1206–1218.

Szmerekovsky JG (2005) The impact of contractor behavior on the client's payment-scheduling problem. *Management Sci.* 51(4):629–640.

Trietsch D, Baker KR (2012) PERT 21: Fitting PERT/CPM for use in the 21st century. *Internat. J. Project Management.* 30(4):490–502.

Van Horne J (1966) Capital-budgeting decisions involving combinations of risky investments. *Management Sci.* 3(2):B84–B92.

Van Slyke RM (1963) Monte Carlo methods and the PERT problem. *Oper. Res.* 11(5):839–860.

Vanhoucke M, Demeulemeester E, Herroelen W (2001) On maximizing the net present value of a project under renewable resource constraints. *Management Sci.* 47(8):1113–1121.

Wiesemann W, Kuhn D (2015) The stochastic time-constrained net present value problem. Schwindt C, Zimmermann J, eds. *Handbook on Project Management and Scheduling Vol. 2.* (Springer, New York), 753–780.

Yan X, Xu J, Zhu Y, Wang J, Yang Y, Wang CX (2016) Downlink average rate and SINR distribution in cellular networks. *IEEE Trans. Commun.* 64(2):847–862.

FACULTY OF ECONOMICS AND BUSINESS
Naamsestraat 69 bus 3500
3000 LEUVEN, BELGIË
tel. + 32 16 32 66 12
fax + 32 16 32 67 91
info@econ.kuleuven.be
www.econ.kuleuven.be